SELF-ORGANIZATION, EMERGENT PROPERTIES AND THE UNITY OF THE WORLD

Gerhard Roth[1] *and Helmut Schwegler*[2]

1. *Unity of the world and unity of science*

Natural philosophers and scientists have always struggled for the unity of the world believing that there is only one reality as the ultimate framework for the infinity of phenomena that can be observed. However, up to date there is no accepted unified view of this world, and it is unlikely that there will be one in the near future. Instead of one all-explaining science or theory, there are many scientific disciplines, and most of them have little, if anything, in common. But even worse, within one and the same discipline, the logical and semantical connections between different areas are not always clear, e.g. in physics between the microscopic and the macroscopic domains, in biosciences between molecular biology and organismal biology, in social sciences between sociology and economics. Each of these disciplines and sub-disciplines seems to have phenomena, concepts and laws (if there are laws recognized at all) of its own, irreducible to and unconnected with those of other disciplines and areas. At the end it seems that there is no one world, but as many worlds as disciplines and sub-disciplines of sciences exist, let alone all those worlds of non-scientific experience.

The reasons for this situation are unclear and may be of different kinds. It may well be that many gaps between domains of knowledge can be bridged by means of further development and refinement of theories and concepts and increased amount of empirical data. History of science is filled with examples of bridging previously disconnected areas of knowledge (e.g., optics and electrodynamics). But we cannot exclude that there are gaps so fundamental in their nature that they cannot be bridged regardless of how much empirical knowledge and theoretical insights we might accumulate. We will call these ontological gaps in the sense that they separate domains that contain completely independent phenomena and concepts. As a consequence, theories from ontologically different domains can neither be reduced one to another, nor can they be unified within the framework of a more comprehensive theory.

Recently, there have been renewed attempts to reach a unified view of the world on the basis of concepts such as "self-organization" (e.g, Jantsch 1980, 1982) and "autopoiesis" (Maturana 1982). These concepts, originating (partly) from biosciences, have quickly spread into other disciplines, especially into the social sciences and the humanities. Here, self-organization and autopoiesis became a sort of magic concept. Another recently much discussed concept to deal with the problem of diversity and unity in nature is emergentism. This concept has become extremely popular among scientists from all disciplines, particularly from biology (Mayr 1982, 1984, 1985).

In this paper, we will discuss the questions whether a unified concept of the world is possible and whether ontological gaps exist in the world that would make such a view impossible. We will discuss the problems mainly with respect to biology, because biological systems are systems which on the one hand are strongly bound to physics and chemistry, but at the same time extend into the domains of psychology and social sciences, at least with respect to human beings. Thus, biology is central to any discussion about the unity of the world.

There are two eminent candidates for ontological gaps: the difference between the living and the nonliving world, and the difference between the material and the mental world. Accordingly, we will deal with two questions: (1) Is biology reducible to physics (and chemistry)? (2) Is mind reducible to brain processes? We will discuss the basic epistemological and methodological problems related to these questions, especially in the light of the concepts of self-organization and of emergent properties.

2. *Can biology be reduced to physics?*

There are two prevailing alternative answers to this question. One is physicalist reductionism claiming that biology can be reduced to physics in the sense biological phenomena can - at least in principle - be described in physical terms and explained by laws of physics and that no essentially biological laws exist (cf. Ayala and Dobzhansky 1974). Accordingly, biology has to be regarded either as a sub-discipline of physics like electroynamics or solid-state physics, or treated, like chemistry, as a separate discipline for historical or practical reasons in which, however, all concepts and modes of descriptions can be expressed and explained in physical terms.

The other view is emergentism claiming that biology is autonomous with respect to physics in the sense that there is a large body of biological phenomena that cannot or not entirely be described in physical terms but need additional biological concepts, and that biological laws exist that are fundamentally different from physical laws. These properties and laws, mainly "higher-level" or organismic ones, are called "emergent", because they are "unexpected" and cannot be formulated and explained on the basis of even the most complete knowledge of the physical properties of the components of the system under consideration (Mayr 1982, 1985).

The strongest argument in favour of the reductionist view is that with the great success of biochemistry, molecular and cellular biology there are major parts of biology where a description and explanation of biological processes in chemical and hence physical terms have been successfully achieved. This is true, for example, for metabolic processes inside the cell, for the processes of genome replicaion, transcription and translation, and for many processes at the organ level. Where, in this context, "gaps" exist, there is hope that these will be bridged as a function of progress in biochemistry and molecular and cellular biology.

There are "higher-level" fields of biology, for example developmental biology or neurobiology, where such a "reduction" is very far from being complete. How the fertilized egg (zygote), through cell division, differentiation and migration, develops into the adult organism, is not understood; how the incredible complexity of the brain is produced, how neural networks function in the context of perception, memory and control of behavior is not understood either. But whenever processes have been studied for a sufficient time and in sufficient depth within these fields, they turn out to be - at least in principle - explicable in chemical and physical terms. The establishment of the highly specific connectivity among nerve cells in the brain ("axon pathfinding"), once thought to be completely inaccessible to a physical approach, is close to its explanation as a highly complex interplay of chemical "guiding factors" (e.g. surface glycoprotein molecules) and response mechanisms inside the connecting neuron.

In certain fields of biology, for example evolutionary biology, however, a physicalist reduction seems to be completely out of reach. For the process of evolution, a description seems to need concepts and an explanation seems to need rules that cannot be borrowed from chemistry and physics. This seems to be especially true in the light of the paradox that the components of which organisms are built and which are accessible to a chemical and physical description, are highly uniform among even distantly related groups, but that the organisms at the same time show a high diversity of morphology, physiology, behavior and ecology. This diversity - so the argument goes - cannot result from the components but only from variation in composition which is governed by truly evolutionary, historical laws completely beyond physics and chemistry. The reductionist counterargument states that if one accepts that evolution with all its phenomena results from the interay between the organism (including all genetic and epigenetic processes) and the environment, both of which are processes and phenomena explicable in physical terms, then evolution should be explicable in those terms, too. There is no evidence in evolution for something to come into play even in evolutionary biology that is "in priniple" inaccessible to physics. The complexity of organismal evolution could simply result from the complexity of the initial and boundary conditions of each single evolutionary process which, as a single process, is - at least in principle - describable in physical terms.

Before we further discuss this reductionist view, let us have a look at emergentism which today probably is the most popular scientific concept among biologists. According to Ernst Mayr, one of the leading proponents of emergentism, the characteristics of living organisms "cannot (not even in theory) be deduced from the most complete knowledge of the components, taken separately or in combination... When such systems are assembled from their components, new chaacteristics of the new whole emerge that could not have been predicted from a knowledge of the components" (Mayr, 1985). It is important to see that emergenists like Ernst Mayr do not want to be misunderstood as vitalists (in the sense that organisms contain some mystical "life-making" substance or principle). Mayr accepts that organisms consist of nothing but physico-chemical components. Yet, he sees that the organization of living systems is fundamentally different from that of non-living ones.

Emergent properties are not considered to be characteristic of biological systems only; they are believed to be quite universal and to occur in inanimate systems, too, but are considered especially important in biological systems.

In the reductionist as well as the emergentist view, the main problem consists in the question: To what degree can the properties of complex systems be explained on the basis of the properties of their components? Emergentism, as mentioned above, accepts that organisms consist of nothing but physicochemical components. However, for emergentism the knowledge of the properties of these components is not sufficient to deduce the properties of the system. These result from the organization of the system, including the existence of "hierarchical levels".

Organization of a system is usually understood as the modes of interaction of its components. Accordingly, if a system has certain properties, these result from the modes of interaction of the components. If we accept such a definition, then the basic question is how the properties of the components relate to their modes of interaction. If emergentism wants to be conclusive, it has to assume that there is no or no strict relationship between the two.

This view, however, neglects the simple fact that there is no way of distinguishing between "properties" and "modes of interaction", because properties are defined through modes of interaction, and modes of interaction are defined through properties. In other words: no phenomenon or process has properties per se, indepenent from its interactions with other phenomena and processes. We can describe phenomena and processes only in relation to other phenomena and processes. Thus, every property results from modes of interaction, and every mode of interaction is a property.

This becomes evident when we briefly discuss a standard example of emergent properties, the properties of sodium chloride, table salt, which consists of two chemically aggressive components, sodium and chlorine, but is harmless as a compound. A molecule (or atom) like sodium has a certain physical structure (nuclei and electron shells). This physical structure (e.g. the wave function of electrons) describes nothing but the modes of interaction of the molecule or atom which can be realized in certain experiments (i.e. interactions with certain media). Thus, a phenomenon which under all possible circumstances behaves (interacts with any kind of environment) like a sodium atom is a sodium atom. It makes no sense to claim that a sodium atom has certain properties by itself regardless of its modes of interaction.

Accordingly, properties of a phenomenon necessarily change whenever the modes of interaction change. These changes can be slight or profound; they may remain within the range of our expectation or may be surprising. If we could test a substance under all possible conditions of interaction with other substances, we would be able to predict any of those changes. However, this is not possible, and as a consequence we will often be confronted with "completely new" or "emergent" properties. The interaction of chemical substances can be such that the atomic structure of the two substances is completely altered (by fusion of the electron shells), as is the case with the sodium chloride molecule. Sodium chloride, in this sense, does not "consist" of sodium and chlorine atoms anymore, and it is, thus, not surprising that it has "completely new" properties. However, these new properies result from the properties of the components, viz., the capacity to drastically change their atomic structure in a given way. Therefore, in this popular case of "emergent properties", it is true that the properties of the sodium chloride molecule are by no means detectable at the level of the components; this is impossible because the molecule does not really consist of the elements sodium and chlorine, but of transformed states of these atoms. The same is true if we proceed to more complicated molecules, e.g., biomolecules, and much more complicated systems such as cells and multicellular organisms.

Emergentism makes an inacceptable distinction between properties and modes of interaction. Accepting the fact that properties and modes of interaction are mutually defined, then emergentism becomes a truism because all properties are emergent properties.

Reductionism, in its common version (which is different from quantum chemical reductionism, see below), denies this fact, namely that due to a specific mode of interaction inside a given system, components can drastically change their properties such that they are not identifiable any more with the components prior to the composition of the system or with substances outside the system. This is especially true for living systems, because they have two peculiar properties: they consist of unusual components, macromolecules of a very complex nature, and show unusual properties, those of self-production and self-maintenance (cf. an der Heiden, Roth and Schwegler 1985, 1986). The most important macromolecules of which organisms are built are amino acids, nucleotides, fatty acids and polysaccharides. Amino acids are the building blocks of proteins which, as structural proteins, constitute most of the structure of organisms, and, as enzymes, regulate all physiological processes. Nucleotides, in the form of nucleic acids (DNA, RNA), are the components of the genome which serves as a "manual" for the enzymes to produce other enzymes and to control the physiological processes; as adenosine phosphates (e.g. ATP) they are the most important energy carriers and play an important role in the intracellular signalling system. Fatty acids (e.g. phospholipids) are the basic components of membranes inside and surrounding the cells and have also energy storage functions. Polysaccharides ("sugars") are the primary energy source for organisms.

All these are molecules that are not naturally found outside organisms. Their most important property is that they show the capability of self-organization or self-aggregation in the sense that the complex arrangement is reached "spontaneously" whenever the appropriate building blocks are present. This is not only true for the three-dimensional folding of enzymes, but also for more complicated structures such as enzyme complexes, ribosomes, membranes and viruses. Since many of these phenomena of self-organization and self-aggregation occur even in vitro, they must be caused by inherent properties of the macromolecules themselves. However, such first-order self-organizing processes - although being the necessary basis of life - have only a limited capacity for the generation of order. Finally, when forming a closed circle of mutual production and maintenance, they form autopoiesis which, thus, can be understood as a circular concatenation of first- and second-order self-organizing processes (Roth 1986).

For a reductionist, autopoiesis, as the typical organization of life, is the direct result of the self-organizing properties of the components, because only from them result the modes of interaction that lead to autopoiesis. In this sense, life and all its properties can be reduced to the properties of the building blocks. However, at present it would be impossible to explain the properties of these biomolecules only on the basis of the properties of their components, for example predict the enzymatic activity of a protein from the amino acids of which it is built, if we had no substantial knowledge about how living systems are organized. So, reductionism in biology is always an ex-post reductionism: we can "deduce" the properties of the system from its components only after we have gained sufficient knowledge about the system as a whole.

The reductionism described so far holds that it can reduce the autopoietic organization of life to abiotic building blocks, the macromolecules, these to smaller molecules and these to atoms, where the sequence of reduction stops. There is another and more sophisticated reductionism that does not rely on such a hierarchy but claims that everything has to be reduced to nuclei and electrons. We will call this quantum-chemical reductionism. This type of reductionism is fully aware of the above described situation in chemical reactions (e.g., regarding our example of sodium chloride) and it does not reduce processes to a level of some objective components and their interactions, because quantum objects like electrons and nuclei adopt the appearance of reality only through interaction with (or measurement by) the observer. Therefore, reduction occurs by a construction of the properties of all (as the radical advocates of quantum-chemical reductionism believe) higher-level objects and processes from quantum theory.

However, even for the simplest molecules, this program can be carried out only by the use of drastic approximation and ad-hoc assumptions. The more complicated the molecules are, the more ad-hoc assumptions have to be applied that are guided by empirical (i.e. post-hoc) knowledge about the properties and the reactions of the molecules under consideration. Therefore, particularly for the large and complex biomolecules, a rigorous quantum-theoretical calculation "from first principles" or "ab initio" turns out to be a fiction, let alone a quantum-chemistry of organismal processes.

Thus, without any possibility to confirm or falsify this quantum-chemical fiction, there is no alternative to introducing the properties of those highly organized systems by their empirically determined modes of interactions which have to be studied under the conditions of life (or under conditions of which we know that they are very similar to those of life). Even the most detailed quantum-chemical knowledge about amino acids would not tell us what decisive role proteins (as sequences of amino acids) could possibly play when interacting with nucleic acids, phospholipids, polysaccharides and many more highly "improbable" chemical substances. We need at least partial knowledge about the organization of life in order to understand how life works. But this is not reductionism anymore, because it says nothing other than life originates from the interaction of specific building blocks under conditions that are typical for life.

Thus, to be able to understand, i.e. to explain, how life originates from its components, is not the same as to reduce life to its components. Rather, it is the opposite of reductionism in that we understand the properties of the components after we have understood their role inside organisms.

What we want to propose is a distinction between reductionism and physicalism. Physicalism is understood here as the attempt to describe all phenomena in terms of physics. The description of observable phenomena (even in atomic and nuclear physics and even more in biology) is possible on the basis of classical physics. In this sense, we can formulate a non-reductionist physicalism with respect to biology: all biological phenomena can be described in terms of classical physics. For explanations, low-level as well as high-level laws are allowed even if (as is usually the case) higher-level laws are not reduced to (explained by) lower ones. Also, "biological laws" are permitted, but, of course, no laws should contradict each other. This physicalism includes any kind of "emergent" properties of complex systems, i.e. system properties that follow "higher-level laws" as compared to the properties of the components of the systems. What is excluded are "ontological" gaps between the different levels, because the whole body of "reality" (see below) is described by the same language of classical physics.

3. *Is mind reducible to the brain?*

The relationship between mind and body, or in more modern terms, between mind and brain, has seen an almost sensational rebirth in recent years. This is not quite surprising, because the stormy progress of brain science and the equally stormy development of computer sciences and artificial-intelligence research has forced theoreticians of any kind to thoroughly reconsider the mind-brain problem (Gardner 1985; Churchland 1986).

It is not our task to present and discuss the numberless theories related to the mind-brain relationship. Many of them have become rather obsolete, viz., all those theories that see mind as being an independent entity with respect to the brain and the body. Empirical evidence shows that mind and mental capacities are strictly dependent on the presence of a brain and change with changes inside the brain. If we accept this, then the only question worth discussing is: How close is the relationship between mind and brain?

A now widely accepted view among philosophers and brain scientists is that mind and brain are identical and that they represent only two aspects of one and the same entity, an "external" (the material brain) and an "internal" one (conscious experience) (Feigl 1967). Indeed, there seem to be good reasons to adhere to such a view. Modern experimental brain science and clinical neurology demonstrate a close correspondence between brain processes and mental processes. Since structure and functional organization of certain subsystems of the mammalian/human brain, for example the visual system or the motor system, are now well known, it is possible to correlate a lesion or dysfunction in a circumscribed brain area with a loss or impairment in visual perception, imagery, memory, or motor control. In some cases, this correspondence can be traced down to the level of single nerve cells. With the refinement of modern scanning techniques (e.g. a combination of PET-scan tomography and NMR technique, cf. Raichle 1986) it may soon be possible to "read" thoughts, i.e. to tell the actual mental activity of a subject, at least in some rough manner, from brain activity. Does this mean that mental activity, mind, can be indeed reduced to neural activity?

Our view is that even if modern brain science can demonstrate an arbitrarily close correspondence between neural states and mental states, this does not allow us to state an identity between both events. First and most trivially, an arbitrarily close correspondence between two phenomena A and B does not necessarily mean an identity of A and B. B can simply be a consequence of A. Streets become wet when it rains, but no one would consider rain and wetness of the street to be identical phenomena. Likewise, we could say that mind is a function or consequence of brain events without implying that mind is identical with the brain.

Furthermore, not all brain processes are accompanied by mind (in the sense of conscious experience). As far as we know, this is true only for cortical processes, and even not for all of them. This means that neural events do not necessarily lead to mind, and that mind is not a consequence of the mere complexity of neural networks (as was often assumed). Cortical processes, in order to be mind-related, apparently need, in addition to sensory inputs, the highly dynamic interaction with several other subsystems, e.g. centers in the brainstem reticular formation that control wakefulness and attention, and with the memory-access system in the limbic system. Thus, there are no single parts and centers of the brain with which we could identify mind. Rather there is a very specific mode of interaction between parts of the brain that leads to or is accompanied by mind.

Finally, modern clinical scanning and EEG techniques have shown that the close correspondence between mind and brain may be realized differently in different individual brains, at least as regards "higher" cognitive functions, such as language and music recognition, planning and imagery; only in one and the same individual, is there a high degree of stereotyped correlation between these mental states and the activity of specific brain areas. Particularly the studies on the (asymmetric) distribution of cognitive functions in the cortex have revealed a high degree of inter-individual variability. While basic perceptive functions usually occur in the same brain areas in all human beings, due to the basic, experience-independent topography of the brain, with respect to "higher" cognitive functions the developing and self-organizing brain-mind system has a higher degree of freedom to specify which parts of the cortex will be involved.

This plasticity is possible only because neural activity is, as such, semanically neutral: a particular neural activity can achieve arbitrary meaning, depending on the context in the brain. This context is specified (i) by the topography of the brain area under consideration, i.e. the set of inputs and outputs of this area, and (ii) by the spatio-temporal processes within the larger networks in which the specific area is situated. Most non-cortical parts of the brain are highly complex in their cytoarchitecture and rigid in their processing capacities, and this means a low degree of variability in the relationship between structure and function. In contrast, the cortex is architecturally highly uniform, and this uniformity means a high degree of versatility with regard to the correlation between structure and function (Creutzfeldt 1983; Rakic and Singer 1988). This constitutes meaning of the cortical processes.

Therefore, the strict correlation between complex cortical processes and their meaning is not pre-established, but is formed during development of the cognitive system through self-organization and self-specification. Initial neural networks develop a certain function that has certain consequences for other neural networks including motor areas that control behavior. These consequences are what we subjectively experience as meaning. They influence the further development of neural networks, and this again leads to modified functions which, through their behavioral or nonbehavioral (brain-internal) consequences, organize new networks etc., etc. Thus, the relationship between brain and function/meaning within the cognitive system develops in an interactive and self-organizing manner during the development of the whole system.

A good example for the relationship between mind and brain is written language. There is no pre-established correlation between the word as a sequence of characters (and of the physical shape and nature of characters) and its meaning. Any word can, in principle, have any kind of meaning, and many words have indeed different meaning at the same time in different contexts or at different historical times. However, at a given time and in a given context, every word must have precisely one meaning, otherwise communication would be impossible. Thus, although word and meaning are by no means identical and can change in any direction, at a given time and in a given context we find a reliable correlation. As in the brain, this correlation between word and meaning is a product of self-organizing and self-specifying processes.

Mind is an emergent property of the brain as being a complex, self-organizing system. As such, mind cannot be reduced to properties of single nerve cells or networks, because what the activity of cells and networks means to the whole brain is asymptotically determined by the interaction of the whole brain, especially with respect to behavioral consequences. However, despite its non-reducibility, the emergence of mind depends on very specific properties and modes of interactions of the components, nerve cells and neural assemblies, with all their electrochemical and dynamical properties, and on a very specific interaction of certain parts of the brain, e.g., cortex, brainstem, limbic system.

Thus, on the one hand, mental processes cannot be reduced to brain processes, because they are agents that - once arisen - are able to modify brain processes which in turn lead to the modification of mental processes. On the other hand, we have seen that all mental processes at a given time have exactly one neural equivalent. This results in the possibility that mind can be described in neural, i.e. physical terms, even if mind is not directly accessible to a physical description. If we really understood all the modes of interaction between the neural networks and all their initial and boundary conditions, we could, at least in principle, derive from the neural processes and their changes the activity of mind. In practice, however, this will never be possible, because by finding out the modes of interactions of neural networks and their initial and boundary conditions we would necessarily interfere with them and change them. What remains, however, is the physical description of the brain. With respect to the mind-brain relationship, as in the case of the biology-physics relationship, we end up with a non-reductionist physicalism.

4. *Reality and Actuality*

There is a fundamental reason why mind is not reducible to brain events, and this has to do with the fact that we try to describe and explain how mind results from the brain while being mind: the system that describes is identical with the system that is being described. Such a circular situation usually leads to a paradox.

In order to cope with this problem, let us, for a moment, imagine what God sees while looking at the world. God sees a world, which we will call REALITY. In this world organisms exist that have to survive within their respective environments. Some of these organisms possess brains that generate and control behavior in such a way that the organisms are able to survive.

At least some of these brains have conscious experience or mind, that forms a world in itself, the phenomenal world or ACTUALITY ("Wirklichkeit"). An "Ego" exists that experiences itself as the subject of this phenomenal world by interaction with other Egos, produces the world of social events and processes.

Let us turn now to a subjectivistic scenario and consider the world in the way in which this Ego experiences its world, or, as we experience our world. There is the environment, filled with objects and processes of an infinity of qualities and quantities; there is our body and its various sensations: touch, position sense, pleasure, pain; and there is the domain of our emotions and thoughts, or "mental" states. These domains are experienced as being different from each other, although they are continuous in a certain sense: the environment and our body seem to belong to the same world, the physical world, which we encounter via our sense organs, although the sensation of our body is different from that of environmental physical objects and processes, and we are constantly aware of our body via a system called proprioception even if we do "nothing specific". The body is felt as being "embedded" in the environment. Our emotions and thoughts seem very different from the phenomena of the environment, but they are intimately embedded in our body: our happiness, fear, anger and pleasure are able to affect our body completely; our thoughts can conquer our brain completely and bodily pain can deeply disturb our mental states.

A paradox arises when we try to combine these two views. Within the objectivistic scenario it seems clear that the phenomenal world, ACTUALITY, is part of, or produced by, the objective world, REALITY, via the nervous system and its functioning. In this sense, people say that our perception, cognition and feeling are produced by or arise "inside" our head or brain. But at the same time, it is said that it is the brain that constitutes all this world we experience! How can these two views be harmonized?

Let us suppose, I am at this moment a patient undergoing brain surgery. I am lying in the operating theater, with my skull open and my brain exposed. I can hear and see what is going on around me, because brain surgery is usually carried out without anesthesia in order to test the site of important brain functions prior to removing parts of the brain (e.g. those infected by a tumor). I can see and feel my body, and I can, with the help of a mirror or a video camera, see my own brain. This brain is, with no doubt, part of the environment, because I can look at it. At the same time, I am, as a neurobiologist, forced to assume that my brain is that system that produces everything I can see, hear, feel and even think. How, then, could this brain look at itself? How could it be outside itself? The whole operating theater, the persons surrounding me, my own body, even my brain should be inside my brain! Since the brain at which I am looking is part of the operating theater which is inside my brain, this would mean that my own brain were inside my own brain. This certainly is a deep paradox.

One could argue that this is a rather unlikely situation. However, it is simply a drastic illustration of the fundamental epistemological and ontological problems that arise whenever we try to develop a theory of cognition that is epistemologically and empirically well founded. This becomes evident, for example, in Maturana's theory as presented in his "Biology of Cognition" (Maturana 1970, 1982). What Maturana was originally interested in was to answer the question: "How do we cognize and know?". He realized that cognition could not be explained from itself, but only as a biological phenomenon; thus, in order to understand cognition he had to understand how living systems are organized. This led him to the question: How does the organization of the living beings condition cognition in general and self-cognition in particular?. But in order to answer this question he had to analyze living systems, to be an observer. One of the central statements in the "Biology of Cognition" is that everything that is said is said by an observer. But this observer - as Maturana emphasizes - is himself a living system, and every explanation of cognition as a biological phenomenon has to contain an explanation of the observer and his role therein. In other words: the observer observes and explains cognition by observing and explaining living systems while being a living system and while exerting cognition.

Accordingly, in Maturana's theory we find a peculiar concatenation of three different types of theory: epistemology (observer), cognitive theory (cognition) and biological systems theory (autopoiesis). By linking these theories in an explanatory circle, Maturana hoped to come to a self-explanatory theory. He started to explain living systems as autopoietic systems that possess brains which, through self-interaction, produce cognition, and, through self-description or language, develop the observer. This observer, then, is able to develop a theory of living systems as autopoietic systems that possess brains which, through self-interaction, produce cognition, and, through self-description or language, develop the observer who, then, is able to develop a theory of living systems etc., etc. If this circle really worked, it would yield the first complete theory in the history of philosophy and science, i.e. a theory that is able to explain its own logical, epistemological and empirical fundaments. And certainly, if it is true that both cognition and the observer are biological phenomena and that they are understood if and as soon as the autopoietic organization of living systems is fully understood, then we have indeed bridged the ontological gap between observer (mental states) and living system (physical states).

But how does one begin a description of this circular system of relations (in circular systems, beginning is always problematic)? Should we take the biological existence of the observer for granted and start with epistemological considerations about what the observer does, how he distinguishes objects from a background, identifies relations and properties, and finally specifies the organization of living beings, and himself as a living being? But then we would end up with the painful question about the ontological status of the statements about the organization of organisms. If they are observer-dependent, how can they answer the question about the "real" origin of the observer without running into a vicious circle? If they are observer-independent, how can an observer observe them at all, given the semantic closure of the cognitive system of the observer that is always emphazised by Maturana? Or should we, in contrast, take the epistemological conditions of the observer for granted and start with biological, even biophysical and biochemical investigations about how organisms are organized, how they interact with the environment, develop a brain as part of the autopoietic network that exerts cognition and, through self-description and language, generate the observer? But what epistemological and ontological status has, then, the description of the organization of living beings? Are they statements about the "real" world? If this were the case, then the observer had access to this world, and this would again contradict the concept of semantic closure of the cognitive/observing system. But if they were not statements about the real world, how could we demonstrate the biological nature of the observer?

The true dilemma arises if we take both the biological and the epistemological status of the observer for granted and concentrate on the investigation of the nature of cognition. Do we, then, talk about cognition in an "objective" manner as resulting from the "real" organization of living organisms (as Maturana proposed) or in an observer-dependent manner? Is cognition a process belonging to the organism or to the observer? Does the observer result from the cognition of the organism? If yes, then the observer has to observe himself as emerging from the "real" organism. But if cognition results as an act of the observer, then we can say nothing about the emergence of cognition and of the observer from the "real" organism.

It becomes clear that by using the ingenious circular relationship between organism, cognition and observer as proposed by Maturana, we run again and again into fundamental onto-epistemological problems. The idea of the circular relationship (or circular constitution) between organism, cognition and observer is a fatal idea because it is based on an ontological jump between REALITY and ACTUALITY, between the real biological system that has a brain and produces the observer, and the actual observer who is a cognitive system and describes biological systems in his phenomenal world. Those biological systems that are assumed to have brains that produce the observer do not exist within the ontological domain, ACTUALITY, in which the observer exists. They exist in the REALITY which is completely inaccessible to any observer (even though we can talk about this REALITY within ACTUALITY). In contrast, those organisms and brains which the observer can study and describe are not those organisms and brains that produce the observer. When I, as the above patient with an open skull, look at "my" brain, then this brain is, of course, not that brain that produces my phenomenal world, but it is a brain that is part of my phenomenal world. The "real" brain that is assumed to produce this phenomenal world including my own Ego does not appear in this phenomenal world.

The strongly seductive idea to establish a self-explanatory system by combining a theory of biological systems, a theory of cognition and a theory of the observer, led Maturana to become an objectivist despite all his attempts to demonstrate the semantic closure of the nervous system. This becomes evident in his writings when he specifies the "true", i.e. observer-independent, nature of organisms (e.g., the statement that organisms are mechanistic, purposeless, etc., systems). However, our phenomenal world is not transcendable; it is truly and completely closed, as Maturana rightly says.

Thus, we cannot reduce mind (observer) to matter/brain and explain how mind (observer) arises from matter/brain, because that brain and that matter that produces our mind (us as observers) is inaccessible to us, and that brain and that matter that is accessible is, as a part of ACTUALITY, not the producer of mind. The only thing we can do is to observe inside our phenomenal world the correlation between brain processes and behavioral reactions (including statements about mental states or self-experience in the case of a self-experiment). As mentioned above, this correlation seems to be very close. To what degree this holds for the relationship between the real brain and the mind that constitutes the phenomenal world in which we exist is a question that is unanswerable because the real world is truly and completely inaccessible. There are no words and no concepts by means of which we could describe a truly consciousness-independent world.

5. *Conclusions*

All we can say about the unity of the world and the unity of the description of this world refers to the phenomenal world. Thus, what we have proposed above concerns the unity of ACTUALITY. This world - the only world scientists can deal with - is neither reducible to a few basic phenomena and laws, as reductionism claims, nor does it consist of unbridgeable domains, as emergentism states. We propose a distinction between the possibility of a unified way of describing the phenomena of this world, and the possibility of a unified way of explaining it. There is the fact that at different levels of complexity systems show properties that as such are not reducible to the properties of their components and can be understood as being "emergent" properties. The reasons for such an emergence of properties result from the fact that properties and modes of interaction define each other mutually, and that whenever certain entities become components of a new system they will show new modes of interaction and, with this, new properties. To what degree these properties can be predicted ab initio from the properties of the entities prior to forming the system is a question of the complexity of the system and the amount of pre-knowledge. In many cases, especially with respect to biological systems, such an ab initio prediction is impossible, because the components, biomolecules, reveal their specific properties only or predominantly within the autopoietic network of the organisms. The only thing we could do (although even this is not possible at present) is to explain ex post how the different biomolecules interact in order to constitute life.

A similar situation is found with respect to the relationship between (actual) brain and mind. It is impossible to reduce mind to brain processes, because brain processes as such are not mind, having no meaning as such. What meaning brain processes have, depends on the functional organization of the brain, e.g., its topography, and the modes of interaction between neural networks. So, the whole system (or at least major subsystems of the brain) specifies the correlation between brain processes and meaning on the basis of the consequences of previous specification. What brain scientists can do is to analyze which of the permanent and actual properties of neural networks lead to specific mental states.

The language in which we describe the origin of life from biomolecules and the origin of mind from brain processes is the language of physics, i.e. a language that refers to processes of spatio-temporal interaction of tissues, cells and molecules. If we try to establish a unified description of the phenomenal world, then there is no alternative to this language. This does not mean that we always have to use this language. It might be extremely cumbersome to describe behavior in terms of interaction of molecules. Many different level- and complexity-dependent languages are possible. It only means that all other languages used in biology have to be able to be translated in principle into the physical language. Accordingly, no concepts and terms are allowed that are not translatable into physical terms.

Also, this physicalism does not imply that everything can be reduced to physical laws. We believe that truly biological laws exist, e.g. evolutionary laws, that can be expressed in physical terms without being reducible to physical laws. Our physicalism, thus, is a non-reductionist physicalism (we could also say, a physicalist emergentism, as opposed to a holistic emergentism), because the causes for law-like processes are not necessarily found at lower levels.

We believe in a unity of description of the phenomenal world that includes the acknowledgement of different levels of complexity of systems. The deep reason for the difficulties that many philosophers of science have with an understanding of the complexity of the world, which leads them either to reductionism or to a holistic emergentism, is the adherence to a substantialism that has dominated philosophy and science for more than two thousand years: the belief that objects have a true nature (substantia) and a true identity and possess properties (accidentia) that can change without a change in the substance of the objects. Any profound change of objects, then, either has to be denied (as does reductionism), or has to be considered inexplicable (as does emergentism). Our alternative view is that the properties of the objects and processes are constituted by interaction, which allows any kind of smooth or dramatic change of properties, and that there is no "true" nature of objects beyond these interaction.

University of Bremen
Brain Research Institute[1]
and Department of Physics[2]

LITERATURE CITED

an der Heiden, U., G. Roth, and H. Schwegler (1985): Principles of self-generation and self-maintenance. Acta Biotheoretica 34: 125-138.

an der Heiden, U., G. Roth, and H. Schwegler (1985): Die Organisation der Organismen. Selbstherstellung und Selbsterhaltung. Funkt. Biol. Med. 5: 330-346

Ayala, F.J., and T. Dobzhansky (eds.) (1974): Studies in the Philosophy of Biology. Berkeley and Los Angeles

Churchland, P.S. (1986): Neurophilosophy. MIT Press, Cambridge, Mass.

Creutzfeldt, O. (1983): Cortex Cerebri. Springer-Verlag, Berlin-Heidelberg-New York.

Feigl, H. (1967): The 'mental' and the 'physical'. In: Feigl, H., M. Scriven, and G. Maxwell (Eds.): Minnesota Studies in the Philosophy of Science, Vol. II: Concepts, Theories, and the Mind-Body Problem. Minneapolis, 370-497.

Gardner, H. (1985): The Mind's New Science. Basic Books, New York.

Jantsch, E. (1980): The unifying paradigm behind autopoiesis, dissipative structures, hyper- and ultracycles. In: Zeleny, M. (ed.): Autopoiesis. Dissipative Structures, and Spontaneous Social Orders. Westview Press, Boulder, Co., 81-87.

Jantsch, E. (1982): Die Selbstorganisation des Universums. DTV, München.

Maturana, H.R. (1970): Biology of Cognition. BCL, Urbana, Ill.

Maturana, H.R. (1982): Erkennen: Die Organisation und Verkörperung von Wirklichkeit. Vieweg, Braunschweig.

Mayr, E. (1982): The Growth of Biological Thought. Harvard University Press, Cambridge, Mass.

Mayr, E. (1984): Die Entwicklung der biologischen Gedankenwelt. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo.

Mayr, E. (1985): How biology differs from the physical sciences. In: D.J. Depew and B.H. Weber (eds.): Evolution at a Crossroads: The New Biology and the New Philosophy of Science. MIT Press, Cambridge, Mass.

Raichle, M.E. (1986): Neuroimaging. Trends in NeuroSciences, Oct. 1986, 525-529.

Rakic, P., and W. Singer (1988): Neurobiology of Neocortex. Wiley, Chichester, New York, Brisbane, Toronto, Singapore.

Roth, G. (1986): Selbstorganisation- Selbsterhaltung-Selbstreferentialitt: Prinzipien der Organisation der Lebewesen und ihre Folgen für die Beziehung zwischen Organismus und Umwelt. In: A. Dress, H. Henrichs, and G. Küppers (Eds.): Selbstorganisation. Die Entstehung von Ordnung in Natur und Gesellschaft. Piper, München-Zürich, 149-180.